

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 11-119798

(43)Date of publication of application : 30.04.1999

(51)Int.Cl.

G10L 9/00
G10L 9/16
// H03M 7/40

(21)Application number : 09-285458

(71)Applicant : SONY CORP

(22)Date of filing : 17.10.1997

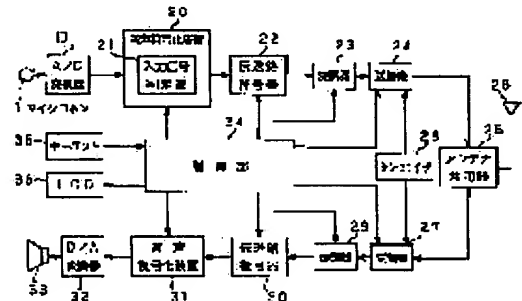
(72)Inventor : IIJIMA KAZUYUKI
NISHIGUCHI MASAYUKI

(54) METHOD OF ENCODING SPEECH AND DEVICE THEREFOR, AND METHOD OF DECODING SPEECH AND DEVICE THEREFOR

(57)Abstract:

PROBLEM TO BE SOLVED: To generate background noise with unnatural feeling relaxed by using encoded data by outputting plural kinds of parameters after a predetermined time in a speech encoding process when a judgment result shows background noise section.

SOLUTION: An input signal judgment device 21 judges whether an input signal supplied from an input terminal 1 and converted into a digital signal by an A/D converter 10, is voice sound V or unvoiced sound UV or background noise in a predetermined time section. And, an idVUV parameter is outputted which shows the result of V/UV judgment such as '0' presenting unvoiced sound, '1' presenting background noise, '2' presenting 1st voiced sound or '3' presenting 2nd second voiced sound. When a voice encoding device 20 grasps from the judgment result in this input judgment part 21, namely, the idVUV parameter that the input signal is a background noise section, the voice encoding device 20 outputs plural kinds of parameters for UV after a predetermined time.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開平11-119798

(43)公開日 平成11年(1999) 4月30日

(51)Int.Cl.⁵

識別記号

F I

G 1 0 L 9/00

G 1 0 L 9/00

J

F

9/16

9/16

// H 0 3 M 7/40

H 0 3 M 7/40

審査請求 未請求 請求項の数13 O L (全 18 頁)

(21)出願番号 特願平9-285458

(22)出願日 平成9年(1997)10月17日

(71)出願人 000002185

ソニー株式会社

東京都品川区北品川6丁目7番35号

(72)発明者 飯島 和幸

東京都品川区北品川6丁目7番35号 ソニ
ー株式会社内

(72)発明者 西口 正之

東京都品川区北品川6丁目7番35号 ソニ
ー株式会社内

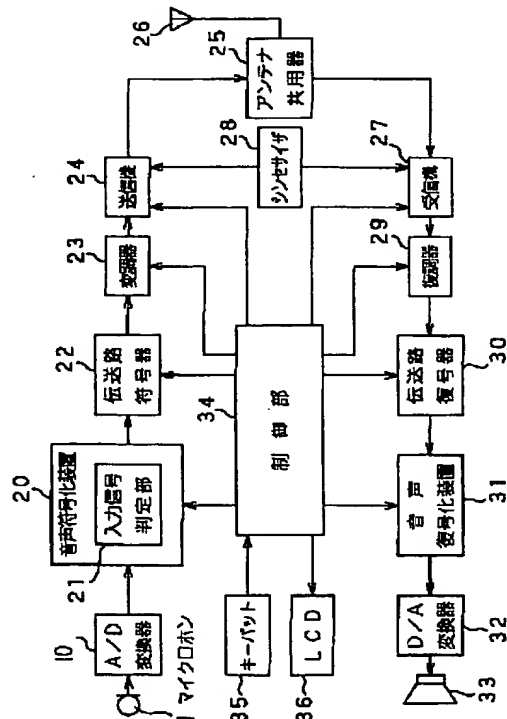
(74)代理人 弁理士 小池 晃 (外2名)

(54)【発明の名称】 音声符号化方法及び装置、並びに音声復号化方法及び装置

(57)【要約】

【課題】 可変レート符号化データを復号化するときには、復号化装置自体で背景雑音を生成するが、過去のパラメータをそのまま繰り返し用いて生成すると、雑音自体がピッチを持つような印象を受け不自然な雑音になることが多い。

【解決手段】 携帯電話装置において音声符号化装置20は、有声音又は無声音区間に分けられる音声信号区間と、背景雑音区間からなる入力信号を、各区間の判定結果に基づいた可変レートで符号化する。この音声符号化装置20は、上記判定結果が上記背景雑音区間を示すときには、所定時間において上記複数種類のパラメータを出力する。なお、上記判定結果を示すパラメータについては、基本パラメータ或いはモードビットとして上記所定時間中にも常に出力する。また、音声復号化装置31は、上記背景雑音区間中には、所定時間において伝送されてきた複数種類のパラメータを用いて上記背景雑音を生成すると共に、上記所定時間中には過去に送られたパラメータを用いて上記背景雑音を生成する。



【特許請求の範囲】

【請求項1】 有声音又は無声音区間に分けられる音声信号区間と、背景雑音区間からなる入力信号を、各区間の判定結果に基づいた可変レートで符号化する音声符号化方法であって、

上記入力信号を時間軸上で所定の符号化単位で区分して各符号化単位で符号化を行って複数種類の音声符号化パラメータを出力する音声符号化工程を備え、

上記音声符号化工程は、上記判定結果が上記背景雑音区間を示すときには、所定時間において上記複数種類のパラメータを出力することを特徴とする音声符号化方法。

【請求項2】 上記音声符号化工程は、上記各区間の判定結果を基本パラメータとして上記所定時間中にも常に出力することを特徴とする請求項1記載の音声符号化方法。

【請求項3】 上記音声符号化工程は、入力信号の短期予測残差を求める短期予測残差算出工程と、求められた短期予測残差をサイン波分析符号化するサイン波分析符号化工程又は上記入力信号を波形符号化により符号化する波形符号化工程とを備えてなることを特徴とする請求項1記載の音声符号化方法。

【請求項4】 上記音声信号区間が有声音であるときには上記サイン波分析符号化工程により、無声音であるときには上記波形符号化工程により上記入力信号を符号化することを特徴とする請求項3記載の音声符号化方法。

【請求項5】 上記音声符号化工程は、上記背景雑音区間中、又は1フレーム前が背景雑音区間であるときには上記短期予測残差算出工程にて差分量子化を行わないことを特徴とする請求項3記載の音声符号化方法。

【請求項6】 有声音又は無声音区間に分けられる音声信号区間と、背景雑音区間からなる入力信号を、各区間の判定結果に基づいた可変レートで符号化する音声符号化装置であって、

上記入力信号を時間軸上で所定の符号化単位で区分して各符号化単位で符号化を行って複数種類の音声符号化パラメータを出力する音声符号化手段を備え、

上記音声符号化手段は、上記判定結果が上記背景雑音区間を示すときには、所定時間において上記複数種類のパラメータを出力することを特徴とする音声符号化装置。

【請求項7】 有声音又は無声音区間に分けられる音声信号区間と、背景雑音区間からなる入力信号を、各区間の判定結果に基づいた可変レートにより符号化して伝送されてきた音声符号化データを復号化する音声復号化方法であって、

上記背景雑音区間中には、所定時間において伝送されてきた複数種類のパラメータを用いて上記背景雑音を生成すると共に、上記所定時間中には過去に送られたパラメータを用いて上記背景雑音を生成することを特徴とする音声復号化方法。

【請求項8】 上記音声符号化データは、入力音声信号

の短期予測残差を求める短期予測残差算出工程と、求められた短期予測残差をサイン波分析符号化するサイン波分析符号化工程又は上記入力音声信号を波形符号化工程により符号化する波形符号化工程とを備えてなる音声符号化工程により生成されていることを特徴とする請求項7記載の音声復号化方法。

【請求項9】 上記音声符号化データは、上記音声信号区間が有声音であるときには上記サイン波分析符号化工程により、無声音であるときには上記波形符号化工程により符号化されていることを特徴とする請求項8記載の音声復号化方法。

【請求項10】 上記所定時間中に背景雑音を生成するのに用いる過去のパラメータは、上記短期予測残差算出工程が算出した少なくとも短期予測符号化係数であることを特徴とする請求項7の音声復号化方法。

【請求項11】 上記所定時間において伝送されてきた複数種類のパラメータ中の上記波形符号化工程からの符号化出力の前の値との差に応じて上記背景雑音を生成することを特徴とする請求項7記載の音声復号化方法。

【請求項12】 上記波形符号化工程からの符号化出力は、短期予測符号化係数に基づいたゲインインデックスであることを特徴とする請求項11記載の音声復号化方法。

【請求項13】 有声音又は無声音区間に分けられる音声信号区間と、背景雑音区間からなる入力信号を、各区間の判定結果に基づいた可変レートにより符号化して伝送されてきた音声符号化データを復号化する音声復号化装置であって、

上記背景雑音区間中には、所定時間において伝送されてきた複数種類のパラメータを用いて上記背景雑音を生成すると共に、上記所定時間中には過去に送られたパラメータを用いて上記背景雑音を生成する音声復号化手段を備えることを特徴とする音声復号化装置。

【発明の詳細な説明】**【0001】**

【発明の属する技術分野】本発明は、有声音又は無声音区間に分けられる音声信号区間と、背景雑音区間からなる入力信号を、各区間の判定結果に基づいた可変レートで符号化する音声符号化方法及び装置に関する。

【0002】また、本発明は、上記音声符号化方法及び装置によって符号化された符号化データを復号化する音声復号化方法及び装置に関する。

【0003】

【従来の技術】近年、伝送路を必要とする通信分野においては、伝送帯域の有効利用を実現するために、伝送しようとする入力信号の種類、例えば有声音又は無声音区間に分けられる音声信号区間と、背景雑音区間のような種類によって、符号化レートを可変してから伝送することが考えられるようになった。

【0004】例えば、背景雑音区間と判断されると、符

号化パラメータを全く送らずに、復号化装置側では、特に背景雑音を生成することをせずに、単にミュートすることが考えられた。

【0005】しかし、これでは通信相手が音声を発していればその音声には背景雑音に乗っているが、音声を発しないときには突然無音になってしまうことになるので不自然な通話となってしまう。

【0006】そのため、可変レートコーデックにおいては、背景雑音区間として判断されると符号化のパラメータのいくつかを送らずに、復号化装置側では過去のパラメータを繰り返し用いて背景雑音を生成するという行っていた。

【0007】

【発明が解決しようとする課題】ところで、上述したように、過去のパラメータをそのまま繰り返し用いると、雑音自体がピッチを持つような印象を受け、不自然な雑音になることが多い。これは、レベルなどを変えても、線スペクトル対(LSP)パラメータが同じである限り起こってしまう。

【0008】他のパラメータを乱数等で変えるようにしても、LSPパラメータが同一であると、不自然な感じを与えてしまう。

【0009】本発明は、上記実情に鑑みてなされたものであり、可変レート符号化を効率良く実現する音声符号化方法及び装置の提供を目的とする。

【0010】また、本発明は、上記実情に鑑みてなされたものであり、可変レート符号化を実現する音声符号化方法及び装置により符号化された符号化データを用いて、背景雑音を不自然感を緩和して生成できる音声復号化方法及び装置の提供を目的とする。

【0011】

【課題を解決するための手段】本発明に係る音声符号化方法は、有声音又は無声音区間に分けられる音声信号区間と、背景雑音区間からなる入力信号を、各区間の判定結果に基づいた可変レートで符号化する音声符号化方法であって、上記課題を解決するために、上記入力信号を時間軸上で所定の符号化単位で区分して各符号化単位で符号化を行って複数種類の音声符号化パラメータを出力する音声符号化工程を備え、上記音声符号化工程は、上記判定結果が上記背景雑音区間を示すときには、所定時間において上記複数種類のパラメータを出力する。

【0012】ここで、上記音声符号化工程は、上記各区間の判定結果を基本パラメータとして上記所定時間中にも常に出力する。

【0013】また、上記音声符号化工程は、入力信号の短期予測残差を求める短期予測残差算出工程と、求められた短期予測残差をサイン波分析符号化するサイン波分析符号化工程又は上記入力信号を波形符号化により符号化する波形符号化工程とを備えてなる。

【0014】また、上記音声信号区間が有声音であると

ときには上記サイン波分析符号化工程により、無声音であるときには上記波形符号化工程により上記入力信号を符号化する。

【0015】また、上記音声符号化工程は、上記背景雑音区間中、又は1フレーム前が背景雑音区間であるときには上記短期予測残差算出工程にて差分量子化を行わない。

【0016】本発明に係る音声符号化装置は、有声音又は無声音区間に分けられる音声信号区間と、背景雑音区間からなる入力信号を、各区間の判定結果に基づいた可変レートで符号化する音声符号化装置であって、上記課題を解決するために、上記入力信号を時間軸上で所定の符号化単位で区分して各符号化単位で符号化を行って複数種類の音声符号化パラメータを出力する音声符号化手段を備え、上記音声符号化手段は、上記判定結果が上記背景雑音区間を示すときには、所定時間において上記複数種類のパラメータを出力する。

【0017】本発明に係る音声復号化方法は、有声音又は無声音区間に分けられる音声信号区間と、背景雑音区間からなる入力信号を、各区間の判定結果に基づいた可変レートにより符号化して伝送されてきた音声符号化データを復号化する音声復号化方法であって、上記課題を解決するために、上記背景雑音区間中には、所定時間において伝送されてきた複数種類のパラメータを用いて上記背景雑音を生成すると共に、上記所定時間中には過去に送られたパラメータを用いて上記背景雑音を生成する。

【0018】このため、過去の例えば線形予測符号化パラメータをそのまま繰り返し使用することがなく、さらに過去に送られた線形予測符号化パラメータと補間しながら使うので、背景雑音の不自然感を緩和できる。

【0019】ここで、上記音声符号化データは、入力音声信号の短期予測残差を求める短期予測残差算出工程と、求められた短期予測残差をサイン波分析符号化するサイン波分析符号化工程又は上記入力音声信号を波形符号化工程により符号化する波形符号化工程とを備えてなる音声符号化工程により生成されている。

【0020】また、上記音声符号化データは、上記音声信号区間が有声音であるときには上記サイン波分析符号化工程により、無声音であるときには上記波形符号化工程により符号化されている。

【0021】また、上記所定時間中に背景雑音を生成するのに用いる過去のパラメータは、上記短期予測残差算出工程が算出した少なくとも短期予測符号化係数であることのできる。

【0022】また、上記所定時間において伝送されてきた複数種類のパラメータ中の上記波形符号化工程からの符号化出力の前の値との差に応じて上記背景雑音を生成する。ここで、上記波形符号化工程からの符号化出力は、短期予測符号化係数に基づいたゲインインデックス

である。

【0023】本発明に係る音声復号化装置は、有声音又は無声音区間に分けられる音声信号区間と、背景雑音区間からなる入力信号を、各区間の判定結果に基づいた可変レートにより符号化して伝送されてきた音声符号化データを復号化する音声復号化装置であって、上記課題を解決するために、上記背景雑音区間中には、所定時間において伝送されてきた複数種類のパラメータを用いて上記背景雑音を生成すると共に、上記所定時間中には過去に送られたパラメータを用いて上記背景雑音を生成する音声復号化手段を備える。

【0024】このため、過去の例えば線形予測符号化パラメータをそのまま繰り返し使用することがなく、さらに過去に送られた線形予測符号化パラメータと補間しながら使うので、背景雑音の不自然感を緩和できる。

【0025】

【発明の実施の形態】以下、本発明に係る音声符号化方法及び装置、並びに音声復号化方法及び装置の実施の形態について説明する。

【0026】この実施の形態は、本発明に係る音声符号化方法及び装置、並びに音声復号化方法及び装置を図1に示すように、音声符号化装置20、並びに音声復号化装置31として用いた携帯電話装置である。

【0027】この携帯電話装置において、本発明に係る音声符号化方法を適用した音声符号化装置20は、有声音(Voiced: V)又は無声音(UnVoiced: UV)区間に分けられる音声信号区間と、背景雑音(Back grand noise: BGN)区間からなる入力信号を、各区間の判定結果に基づいた可変レートで符号化する音声符号化装置であり、上記入力信号を時間軸上で所定の符号化単位で区分して各符号化単位で符号化を行って複数種類の音声符号化パラメータを出力する。

【0028】この音声符号化装置20は、上記判定結果が上記背景雑音区間を示すときには、所定時間において上記複数種類のパラメータを出力する。なお、上記判定結果を示すパラメータについては、基本パラメータ或いはモードビットとして上記所定時間中にも常に出力する。

【0029】ここで、入力信号が有声音(V)又は無声音(UV)区間であるか、又は背景雑音(BGN)区間であるかは、入力信号判定部21が判定する。

【0030】すなわち、この入力信号判定装置21は、入力端子1から供給され、A/D変換器10でデジタル信号に変換された入力信号が、所定時間区間で有声音(V)、又は無声音(UV)であるか、又は背景雑音(BGN)であるかを判定する。そして、上記無声音であることを示す“0”、上記背景雑音であることを示す“1”、第1有声音であることを示す“2”、又は第2有声音であることを示す“3”というV/UV判定の結果を示すidVUVパラメータを出力する。このidVUVパラメ

ータは、上述したように基本パラメータ或いはモードビットとされて、上記所定時間中にも常に出力される。

【0031】この入力信号判定部21での判定結果、すなわちidVUVパラメータにより、入力信号が背景雑音区間であることを音声符号化装置20が把握すると、この音声符号化装置20は、所定時間において、後述するUV用の複数種類のパラメータを出力する。

【0032】また、この携帯電話装置において、本発明に係る音声復号化方法を適用した音声復号化装置31は、有声音又は無声音区間に分けられる音声信号区間と、背景雑音区間からなる入力信号を、各区間の判定結果に基づいた可変レートにより符号化して伝送されてきた音声符号化データを復号化する音声復号化装置であり、上記背景雑音区間中には、所定時間において伝送されてきた複数種類のパラメータを用いて上記背景雑音を生成すると共に、上記所定時間中には過去に送られたパラメータを用いて上記背景雑音を生成する。

【0033】また、この携帯電話装置は、送信時には、マイクロホン1から入力された音声信号を、A/D変換器10によりデジタル信号に変換し、音声符号化装置20により上記idVUVパラメータに基づいた可変レートの符号化を施し、伝送路符号化器22により伝送路の品質が音声品質に影響を受けにくいように符号化した後、変調器23で変調し、送信機24で出力ビットに送信処理を施し、アンテナ共用器25を通して、アンテナ26から送信する。

【0034】また、受信時には、アンテナ26で捉えた電波を、アンテナ共用器25を通じて受信機27で受信し、復調器29で復調し、伝送路復号化器30で伝送路誤りを訂正し、音声復号化装置31で復号し、D/A変換器32でアナログ音声信号に戻して、スピーカ33から出力する。

【0035】また、制御部34は上記各部をコントロールし、シンセサイザ28は送受信周波数を送信機24、及び受信機27に与えている。また、キーパッド35及びLCD表示器36はマンマシンインターフェースに利用される。

【0036】次に、有声音又は無声音区間に分けられる音声信号区間と、背景雑音区間からなる入力信号を、入力信号判定部21の判定結果idVUVパラメータに基づいて、可変レートで符号化する音声符号化装置20について説明する。

【0037】まず、可変レートエンコードを説明しておく。LSP量子化インデクス、及び励起パラメータインデクスは、以下の表1に示すように、idVUV判定パラメータが“1”の背景雑音のときには、1フレーム20msあたり0ビットにしてしまう。idVUV判定パラメータが“0”の無声音、“2”、“3”の有声音のときには、そのまま20msecあたり18ビット、及び20ビットとする。これにより可変レートエンコードが実現できる。

【0038】

【表1】

モード (idVUV)	背景雑音 (1)	無声音 (0)	有声音 (2,3)
V/UV 判定出力	2bit/20msec	2bit/20msec	2bit/20msec
LSP 量子化インデックス	0bit/20msec	18bit/20msec	18bit/20msec
励起パラメータインデックス	0bit/20msec	20bit/20msec	20bit/20msec
Total	2bit/20msec	40bit/20msec	40bit/20msec

【0039】ここで、idVUV判定パラメータの内の、V/UV判定出力となる2ビットは、常に符号化されている。なお、無声音時の励起パラメータインデックスとしては、後述する雑音符号帳のコードブックのシェイプインデックスと、ゲインインデックスが挙げられる。

【0040】この音声符号化装置の構成を図2、図3に示す。図2の音声符号化装置20の基本的な考え方は、入力音声信号の短期予測残差例えばLPC（線形予測符号化）残差を求めてサイン波分析（sinusoidal analysis）符号化、例えばハーモニックコーディング（harmonic coding）を行う第1の符号化部110と、入力音声信号に対して位相伝送を行う波形符号化により符号化する第2の符号化部120とを有し、入力信号の有声音（V：Voiced）の部分の符号化には第1の符号化部110を用い、入力信号の無声音（UV：Unvoiced）の部分の符号化には第2の符号化部120を用いるようにすることである。

【0041】上記第1の符号化部110には、例えばLPC残差をハーモニック符号化やマルチバンド励起（MBE）符号化のようなサイン波分析符号化を行う構成が用いられる。上記第2の符号化部120には、例えば合成による分析法を用いて最適ベクトルのクローズドループサーチによるベクトル量子化を用いた符号励起線形予測（CELP）符号化の構成が用いられる。

【0042】図2の例では、入力端子101に供給された音声信号が、第1の符号化部110のLPC逆フィルタ111及びLPC分析・量子化部113に送られている。LPC分析・量子化部113で得られたLPC係数あるいはいわゆる α パラメータは、LPC逆フィルタ111に送られて、このLPC逆フィルタ111により入力音声信号の線形予測残差（LPC残差）が取り出される。また、LPC分析・量子化部113からは、後述するようにLSP（線スペクトル対）の量子化出力が取り出され、これが出力端子102に送られる。LPC逆フィルタ111からのLPC残差は、サイン波分析符号化部114に送られる。サイン波分析符号化部114では、ピッチ検出やスペクトルエンベロープ振幅計算が行われると共に、上記入力信号判定部21と同一構成の入力信号判定部115により入力信号の上記idVUVパラメータが求められる。サイン波分析符号化部114からのスペクトルエンベロープ振幅データはベクトル量子化部116に送られる。スペクトルエンベロープのベクトル量子化出力としてのベクトル量子化部116からのコー

ドブックインデックスは、スイッチ117を介して出力端子103に送られ、サイン波分析符号化部114からのピッチ出力は、スイッチ118を介して出力端子104に送られる。また、入力信号判定部115からのidVUV判定パラメータ出力は出力端子105に送られると共に、スイッチ117、118及び図3に示すスイッチ119の制御信号に使われる。スイッチ117、118は、上記制御信号により有声音（V）のとき上記インデックス及びピッチを選択して各出力端子103及び104からそれぞれ出力する。

【0043】また、上記ベクトル量子化部116でのベクトル量子化の際には、例えば、周波数軸上の有効帯域1ブロック分の振幅データに対して、ブロック内の最後のデータからブロック内の最初のデータまでの値を補間するようなダミーデータ、又は最後のデータ及び最初のデータを延長するようなダミーデータを最後と最初に適当な数だけ付加してデータ個数を N_F 個に拡大した後、帯域制限型の O_S 倍（例えば8倍）のオーバーサンプリングを施すことにより O_S 倍の個数の振幅データを求め、この O_S 倍の個数（ $(m_{MX}+1) \times O_S$ 個）の振幅データを直線補間してさらに多くの N_M 個（例えば2048個）に拡張し、この N_M 個のデータを間引いて上記一定個数 M （例えば44個）のデータに変換した後、ベクトル量子化している。

【0044】図2の第2の符号化部120は、この例ではCELP（符号励起線形予測）符号化構成を有しており、雑音符号帳121からの出力を、重み付きの合成フィルタ122により合成処理し、得られた重み付き音声信号を減算器123に送り、入力端子101に供給された音声信号を聴覚重み付けフィルタ125を介して得られた音声との誤差を取り出し、この誤差を距離計算回路124に送って距離計算を行い、誤差が最小となるようなベクトルを雑音符号帳121でサーチするような、合成による分析（Analysis by Synthesis）法を用いたクローズドループサーチを用いた時間軸波形のベクトル量子化を行っている。このCELP符号化は、上述したように無声音部分の符号化に用いられており、雑音符号帳121からのUVデータとしてのコードブックインデックスは、上記入力信号判定部115からのidVUV判定パラメータが無声音（UV）のときオンとなるスイッチ127を介して、出力端子107より取り出される。

【0045】また、スイッチ127の制御信号となるidVUV判定パラメータが“1”となり入力信号が背景雑音

信号であると判定したときには、所定時間、例えば8フレーム分の時間において無声音時の複数のパラメータ、例えば雑音符号帳121からのUVデータとしてのシェイプインデックスやゲインインデックスを送る。

【0046】次に、上記図2に示した音声信号符号化装置のより具体的な構成について、図3を参照しながら説明する。なお、図3において、上記図2の各部と対応する部分には同じ指示符号を付している。

【0047】この図3に示された音声信号符号化装置において、入力端子101に供給された音声信号は、ハイパスフィルタ(HPF)109にて不要な帯域の信号を除去するフィルタ処理が施された後、LPC(線形予測符号化)分析・量子化部113のLPC分析回路132と、LPC逆フィルタ回路111とに送られる。

【0048】LPC分析・量子化部113のLPC分析回路132は、入力信号波形の256サンプル程度の長さを1ブロックとしてハミング窓をかけて、自己相関法により線形予測係数、いわゆる α パラメータを求める。データ出力の単位となるフレーミングの間隔は、160サンプル程度とする。サンプリング周波数 f_s が例えば8kHzのとき、1フレーム間隔は160サンプルで20msecとなる。

【0049】LPC分析回路132からの α パラメータは、 $\alpha \rightarrow$ LSP変換回路133に送られて、線スペクトル対(LSP)パラメータに変換される。これは、直接型のフィルタ係数として求めた α パラメータを、例えば10個、すなわち5対のLSPパラメータに変換する。変換は例えばニュートン・ラプソン法等を用いて行う。このLSPパラメータに変換するのは、 α パラメータよりも補間特性に優れているからである。

【0050】 $\alpha \rightarrow$ LSP変換回路133からのLSPパラメータは、LSP量子化器134によりマトリクスあるいはベクトル量子化される。このとき、このLSP量子化器134は、リーキングファクタを有する差分ベクトル量子化(差分VQ)を行っても良い。差分VQの量子化効率マトリクス量子化(MQ)に比べると優れており、また、MQのように量子化誤差が一方のフレームに偏ることがないため、滑らかで異音の少ない音声を得られる。しかし、差分VQでは一度エラーが起こると、その影響がしばらく続くので、差分のリーク量を大目にとるようにしている。しかし、入力信号判定部115で、背景雑音と判断されたときには、LSPは送らないので、LSP量子化器134では、差分量子化の一種である上記差分VQを行わない。

【0051】このLSP量子化器134からの量子化出力、すなわちLSP量子化のインデックスは、スイッチ119によって切り換えられてから、端子102を介して取り出され、また量子化済みのLSPベクトルは、LSP補間回路136に送られる。ここで、スイッチ119は、上記入力信号判定部115からのidVUV判定フラグ

により切り換えが制御され、例えば有声音(V)のときにオンとなる。

【0052】また、LSP補間回路136は、20msec毎に量子化されたLSPのベクトルを補間し、8倍のレートにする。すなわち、2.5msec毎にLSPベクトルが更新されるようにする。これは、残差波形をハーモニック符号化復号化方法により分析合成すると、その合成波形のエンベロープは非常になだらかでスムーズな波形になるため、LPC係数が20msec毎に急激に変化すると異音を発生することがあるからである。すなわち、2.5msec毎にLPC係数が徐々に変化してゆくようにすれば、このような異音の発生を防ぐことができる。

【0053】このような補間が行われた2.5msec毎のLSPベクトルを用いて入力音声の逆フィルタリングを実行するために、LSP $\rightarrow\alpha$ 変換回路137により、LSPパラメータを例えば10次程度の直接型フィルタの係数である α パラメータに変換する。このLSP $\rightarrow\alpha$ 変換回路137からの出力は、上記LPC逆フィルタ回路111に送られ、このLPC逆フィルタ111では、2.5msec毎に更新される α パラメータにより逆フィルタリング処理を行って、滑らかな出力を得るようにしている。このLPC逆フィルタ111からの出力は、サイン波分析符号化部114、具体的には例えばハーモニック符号化回路、の直交変換回路145、例えばDFT(離散フーリエ変換)回路に送られる。

【0054】LPC分析・量子化部113のLPC分析回路132からの α パラメータは、聴覚重み付けフィルタ算出回路139に送られて聴覚重み付けのためのデータが求められ、この重み付けデータが後述する聴覚重み付きのベクトル量子化器116と、第2の符号化部120の聴覚重み付けフィルタ125及び聴覚重み付きの合成フィルタ122とに送られる。

【0055】ハーモニック符号化回路等のサイン波分析符号化部114では、LPC逆フィルタ111からの出力を、ハーモニック符号化の方法で分析する。すなわち、ピッチ検出、各ハーモニクスの振幅 A_m の算出、有声音(V)/無声音(UV)の判別を行い、ピッチによって変化するハーモニクスのエンベロープあるいは振幅 A_m の個数を次元変換して一定数にしている。

【0056】図3に示すサイン波分析符号化部114の具体例においては、一般のハーモニック符号化を想定しているが、特に、MBE(Multiband Excitation: マルチバンド励起)符号化の場合には、同時刻(同じブロックあるいはフレーム内)の周波数軸領域いわゆるバンド毎に有声音(Voiced)部分と無声音(Unvoiced)部分とが存在するという仮定でモデル化することになる。それ以外のハーモニック符号化では、1ブロックあるいはフレーム内の音声の有声音か無声音かの択一的な判定がなされることになる。なお、以下の説明中のフレーム毎の

V/U/Vとは、MBE符号化に適用した場合には全バンドがU/Vのときを当該フレームのU/Vとしている。ここで上記MBEの分析合成手法については、本件出願人が先に提案した特願平4-91422号明細書及び図面に詳細な具体例を開示している。

【0057】図3のサイン波分析符号化部114のオープンループピッチサーチ部141には、上記入力端子101からの入力音声信号が、またゼロクロスカウンタ142には、上記HPF（ハイパスフィルタ）109からの信号がそれぞれ供給されている。サイン波分析符号化部114の直交変換回路145には、LPC逆フィルタ111からのLPC残差あるいは線形予測残差が供給されている。オープンループピッチサーチ部141では、入力信号のLPC残差をとってオープンループによる比較的ラフなピッチのサーチが行われ、抽出された粗ピッチデータは高精度ピッチサーチ146に送られて、後述するようなクロズドループによる高精度のピッチサーチ（ピッチのファインサーチ）が行われる。また、オープンループピッチサーチ部141からは、上記粗ピッチデータと共にLPC残差の自己相関の最大値をパワーで正規化した正規化自己相関最大値 $r(p)$ が取り出され、入力信号判定部115に送られている。

【0058】直交変換回路145では例えばDFT（離散フーリエ変換）等の直交変換処理が施されて、時間軸上のLPC残差が周波数軸上のスペクトル振幅データに変換される。この直交変換回路145からの出力は、高精度ピッチサーチ部146及びスペクトル振幅あるいはエンベロープを評価するためのスペクトル評価部148に送られる。

【0059】高精度（ファイン）ピッチサーチ部146には、オープンループピッチサーチ部141で抽出された比較的ラフな粗ピッチデータと、直交変換部145により例えばDFTされた周波数軸上のデータとが供給されている。この高精度ピッチサーチ部146では、上記粗ピッチデータ値を中心に、0.2〜0.5きざみで±数サンプルずつ振って、最適な小数点付き（フローティング）のファインピッチデータの値へ追い込む。このときのファインサーチの手法として、いわゆる合成による分析（Analysis by Synthesis）法を用い、合成されたパワースペクトルが原音のパワースペクトルに最も近くなるようにピッチを選んでいる。このようなクロズドループによる高精度のピッチサーチ部146からのピッチデータについては、スイッチ118を介して出力端子104に送っている。

【0060】スペクトル評価部148では、LPC残差の直交変換出力としてのスペクトル振幅及びピッチに基づいて各ハーモニクス大きさ及びその集合であるスペクトルエンベロープが評価され、高精度ピッチサーチ部146及び聴覚重み付きのベクトル量子化器116に送られる。

【0061】入力信号判定部115は、オープンループピッチサーチ部141からの正規化自己相関最大値 $r(p)$ と、ゼロクロスカウンタ142からのゼロクロスカウント値とに基づいて、当該フレームが上記有声音であるか、無声音であるか、あるいは背景雑音であるかの判定を行い、上記idVUV判定パラメータを出力する。この入力信号判定部115からの上記idVUV判定パラメータは、出力端子105を介して取り出されると共に、上述したように、スイッチ119及び、スイッチ117、118及び127の切り換え制御信号としても用いられる。

【0062】ところで、スペクトル評価部148の出力部あるいはベクトル量子化器116の入力部には、データ数変換（一種のサンプリングレート変換）部が設けられている。このデータ数変換部は、上記ピッチに応じて周波数軸上での分割帯域数が異なり、データ数が異なることを考慮して、エンベロープの振幅データ $|A_m|$ を一定の個数にするためのものである。すなわち、例えば有効帯域を3400kHzまでとすると、この有効帯域が上記ピッチに応じて、8バンド〜63バンドに分割されることになり、これらの各バンド毎に得られる上記振幅データ $|A_m|$ の個数 $m_{MX}+1$ も8〜63と変化することになる。このためデータ数変換部では、この可変個数 $m_{MX}+1$ の振幅データを一定個数M個、例えば44個、のデータに変換している。

【0063】このスペクトル評価部148の出力部あるいはベクトル量子化器116の入力部に設けられたデータ数変換部からの上記一定個数M個（例えば44個）の振幅データあるいはエンベロープデータが、ベクトル量子化器116により、所定個数、例えば44個のデータ毎にまとめられてベクトルとされ、重み付きベクトル量子化が施される。この重みは、聴覚重み付けフィルタ算出回路139からの出力により与えられる。ベクトル量子化器116からの上記エンベロープのインデックスは、スイッチ117を介して出力端子103より取り出される。なお、上記重み付きベクトル量子化に先だって、所定個数のデータから成るベクトルについて適当なリーク係数を用いたフレーム間差分をとっておくようにしてもよい。

【0064】次に、第2の符号化部120について説明する。第2の符号化部120は、いわゆるCELP（符号励起線形予測）符号化構成を有しており、特に、入力音声信号の無声音部分の符号化のために用いられている。この無声音部分用のCELP符号化構成において、雑音符号帳、いわゆるストキャスティック・コードブック（stochastic code book）121からの代表値出力である無声音のLPC残差に相当するノイズ出力を、ゲイン回路126を介して、聴覚重み付きの合成フィルタ122に送っている。重み付きの合成フィルタ122では、入力されたノイズをLPC合成処理し、得られた重

み付き無声音の信号を減算器123に送っている。減算器123には、上記入力端子101からHPF（ハイパスフィルタ）109を介して供給された音声信号を聴覚重み付けフィルタ125で聴覚重み付けした信号が入力されており、合成フィルタ122からの信号との差分あるいは誤差を取り出している。なお、聴覚重み付けフィルタ125の出力から聴覚重み付け合成フィルタの零入力応答を事前に差し引いておくものとする。この誤差を距離計算回路124に送って距離計算を行い、誤差が最小となるような代表値ベクトルを雑音符号帳121でサーチする。このような合成による分析（Analysis by Synthesis）法を用いたクローズドループサーチを用いた時間軸波形のベクトル量子化を行っている。

【0065】このCELP符号化構成を用いた第2の符号化部120からのUV（無声音）部分用のデータとしては、雑音符号帳121からのコードブックのシェイプインデックスと、ゲイン回路126からのコードブックのゲインインデックスとが取り出される。雑音符号帳121からのUVデータであるシェイプインデックスは、スイッチ127sを介して出力端子107sに送られ、ゲイン回路126のUVデータであるゲインインデックスは、スイッチ127gを介して出力端子107gに送られている。

【0066】ここで、これらのスイッチ127s、127g及び上記スイッチ117、118は、上記入力信号判定部115からのidVUV判定パラメータによりオン／オフ制御され、スイッチ117、118は、現在伝送しようとするフレームの音声信号のidVUV判定パラメータが有声音（V）のときオンとなり、スイッチ127s、127gは、現在伝送しようとするフレームの音声信号が無声音（UV）のときオンとなる。また、idVUV判定パラメータが背景雑音を表すとき、上記スイッチ127s、127gは、8フレーム時間毎にオンとされ、上記シェイプインデックス、ゲインインデックスを出力する。また、上記スイッチ119も8フレーム時間毎にオンとされ、上記UV用のLSPインデックスを出力する。これらが、上述したUV用の複数種類のパラメータである。

【0067】図4には、上記入力信号判定部（図中、入力信号判定装置と記す。）115の詳細な構成を示す。すなわち、この入力信号判定部115は、入力端子1から実効（root mean square、r.m.s）値演算部2を介して入力された入力信号（実効値）の最小レベルを所定時間区間、例えば20msecで検出し、この最小レベルを保持する最小レベル演算部4と、上記実効値演算部2からの入力信号実効値からリファレンスレベルを演算により求めるリファレンスレベル演算部5とを備え、最小レベル演算部4からの最小レベルとリファレンスレベル演算部5からのリファレンスレベルとに基づいて上記入力信号が所定時間区間で有声音（Voice）、又は無声音（UnVoice）であるか、又は背景雑音であるかを判定する。そ

して、上記無声音であることを示す“0”、上記背景雑音であることを示す“1”、第1有声音であることを示す“2”、又は第2有声音であることを示す“3”というV／UV判定の結果を示すidVUVパラメータを出力する。

【0068】また、この入力信号判定装置21は、上記実効値演算部2からの入力信号実効値の所定時間区間に対して仮に有声音（V）／無声音（UV）とを判定するV／UV判定部3を備え、最小レベル演算部4にV／UV判定結果を供給する。最小レベル演算部4は、このV／UV判定結果に基づいて上記最小レベルを演算する。

【0069】また、V／UV判定部3からのV／UV判定結果はパラメータ生成部8にも供給される。このパラメータ生成部8は、出力端子105から上記idVUVパラメータを出力する。

【0070】このパラメータ生成部8は、最小レベル演算部4からの最小レベルとリファレンスレベル演算部5からのリファレンスレベルとを比較する比較部7からの比較結果と、上記V／UV判定部3からのV／UV判定結果に基づいて上記idVUVパラメータを出力する。

【0071】以下、この入力信号判定装置21の動作について説明する。例えば、走行している電車の中では、背景雑音のレベルが高い。そこで、周りの雑音に合わせてしきい値（スレシヨルド）を決定することが望ましい。

【0072】そこで、上記最小レベル演算部4では、適当な所定時間区間で一番小さいレベルを最小レベルとし、その最小レベルを上記所定時間毎に更新していく。

【0073】図5は最小レベル演算部4のアルゴリズムを示すフローチャートである。このフローチャートでは、最小レベルの更新（トラッキング）を、最小レベル候補値cdLevのセットとクリア、及び最小レベルgmlのセットとクリアに大きく分けている。

【0074】先ず、ステップS1で、V／UV判定部3からのV／UV判定結果に基づいて有声音フレームの連続回数vContが4より多い整数となるかを判断する。すなわち、有声音Vと判断されたフレームが4より大きな整数回、つまり5フレーム連続したか否かを判断する。ここで、有声音フレームが5フレーム以上連続している場合は、音声区間に入っていると判断し、ステップS2に進み、最小レベル候補値cdLevをクリアする。このステップS2では、候補値が設定され続けた回数gmlSetStateは0である。一方、ステップS1で、有声音フレームの連続回数vContが4以下であると判断すると、ステップS3に進む。

【0075】ステップS3では、現在の入力信号の実効値演算部2を介した入力レベルlevが最小レベルの最低値MIN_GMLより小さいか否かを判断する。ここでいう、最小レベルの最低値MIN_GMLは、最小レベルgmlが0とな

らないように決定されている。ここで、入力レベルlevが最小レベルの最低値MIN_GMLより小さいと判断すると、ステップS 4で上記最小レベルの最低値MIN_GMLを最小レベルgmlとして設定する。このステップS 4では、候補値が設定され続けた回数gmlSetStateと、最小レベルが設定された後、候補値が設定されていない回数gmlResetStateとは0である。一方、ステップS 3で現在の入力レベルlevは最小レベルの最低値MIN_GML以上であると判断するとステップS 5に進む。

【0076】ステップS 5では、現在の入力レベルlevが最小レベルgmlより小さいか否かを判断する。ここでYESとなれば、ステップS 6に進む。すなわち、このステップS 6は、上記ステップS 3の判断で現在の入力レベルlevが最小レベルの最低値MIN_GML以上であり、さらに上記ステップS 5で現在の入力レベルlevが最小レベルgmlより小さいと判断されたときに、その入力レベルlevを最小レベルgmlとして設定する。一方、このステップS 5で現在の入力レベルlevが最小レベルgml以上であると判断すると、ステップS 7に進む。

【0077】ステップS 7では、現在の入力レベルが充分小さいか、候補値cdLevとの変動が小さいか否かを判断する。このステップS 7では、現在の入力レベルが充分小さいか、候補値cdLevとの変動が小さいかという判断を、status0であるか否かで判断している。status0は、入力レベルlevが100.0以下というように充分小さいか、又は、入力レベルが500.0以下で候補値cdLev*0.70より大きく、かつ候補値cdLev*1.30より小さいというように候補値cdLevとの変動が小さいことを表す状態である。ここで、YESを選択し、現在の入力レベルが充分小さいか、又は候補値cdLevとの変動が小さいと判断すると、候補値cdLevが更新される。一方、NOを選択するとステップS 11に進む。

【0078】ステップS 8では、候補値cdLevが更新され続けた回数gmlSetStateが7回以上であるか否か、すなわち候補値cdLevの更新が7フレーム連続するか否かを判断する。ここで、候補値cdLevの更新が7フレーム以上連続したと判断すれば、ステップS 9に進み、そのときの入力レベルlevを最小レベルgmlとする。候補値cdLevの更新が6フレーム以内であれば、NOとなり、ス

$$\text{refLev} = A \times \max(\text{lev}, \text{refLev}) + (1.0 - A) \times \min(\text{lev}, \text{refLev})$$

この(1)式において、入力端子6から与えるA=0.75としたときの入力レベルlevとリファレンスレベルrefLevとの関係を図6に示す。リファレンスレベルrefLevは、立ち上がりは入力レベルlevと同様に立ち上がるが、立ち下がりでは緩やかに減少していく。このため、上記リファレンスレベルを用いることで、音声信号区間において、瞬間的にたまたまレベルが下がった状態を背景雑音区間として判定してしまうことを防いでいる。このように、リファレンスレベル演算部5は、瞬間的なレ

ベル変動に対してもある程度余裕を持たせるような滑らかなレベルを演算する。

【0079】一方、ステップS 7での判断でNOとなった後に進んだステップS 11では、1フレーム過去の入力レベルprevLevと現在の入力レベルlevの変動が小さいか否かを判断する。この1フレーム過去の入力レベルprevLevと現在の入力レベルlevの変動が小さいか否かの判断は、status1であるか否かの判断で行っている。status1は、現在の入力レベルlevが100.0以下というように充分小さいか、又は現在の入力レベルlevが500.0以下で1フレーム過去の入力レベルprevLev*0.70より大きく、かつ1フレーム過去の入力レベルprevLev*1.30より小さいというように1フレーム過去の入力レベルprevLevとの変動が小さいことを表す状態である。ここで、YESを選択し、現在の入力レベルが充分小さいか、又は1フレーム過去の入力レベルと現在の入力レベルの変動が小さい場合には、ステップS 12に進み、現在の入力レベルlevを最小レベル候補値cdLevに設定する。一方、ステップS 11で現在の入力レベルが小さくないか、又は1フレーム過去の入力レベルと現在の入力レベルの変動が小さくないと判断すると、ステップS 13に進む。

【0080】ステップS 13では、最小レベルが設定された後、候補値が設定されていない回数gmlResetStateが40より大きいと判断する。ここで、NOとなり候補値が設定されていない回数gmlResetStateが40以下であるときには、ステップS 14で最小レベル候補値cdLevをクリアし、予め定めた最小値をセットする。一方ここで、YESとなり、候補値が設定されていない回数gmlResetStateが40回を越えていると判断すると、ステップS 15に進み、最小レベルgmlは最小レベルの最低値MIN_GMLに設定される。

【0081】以上のように最小レベルはある時間保持され、順次更新される。

【0082】次に、リファレンスレベル演算部5の動作について図6を用いて説明する。このリファレンスレベル演算部5は、リファレンスレベルrefLevを次の(1)式で算出する。

【0083】

$$\dots (1)$$

ベル変動に対してもある程度余裕を持たせるような滑らかなレベルを演算する。

【0084】比較部7は、最小レベル演算部4からの最小レベルに所定の定数Bを掛けたB×gmlと上記リファレンスレベル演算部5からのリファレンスレベルrefLevとを比較する。そして、その比較結果は、パラメータ生成部8に送られる。

【0085】パラメータ生成部8は、V/U判定部3での判定結果により、Vと判定されたフレームについて

は、上記(1)式に示したリファレンスレベル refLev が最小レベル gml のB倍より小さいか否かを調べ、小さいときには背景雑音区間と判断する。ただし、過去のV/U/V判断を調べ、Vのフレームが2フレーム以上連続している場合は音声区間が始まっているものとし、背景雑音区間と判断することはない。すなわち、現在のフレームがVと判断されたときは、過去Vフレームが連続しているかを調べ、連続している場合には背景雑音モードに入らない。これは、Vフレームが連続しているときに背景雑音モードに入ると不連続感が生じるためである。

【0086】また、パラメータ生成部8は、V/U/V判定部3からの判定結果がUVと判定されたフレームについては、Vの場合と同様に、リファレンスレベル refLev が最小レベル gml のB倍より小さいかを調べ、この条件を4回満たした場合、背景雑音区間と判断する。すなわち、UVと判断された場合は、4フレーム連続で上記条件を満たした後、背景雑音区間と判断する。

【0087】なお、上記Bは適当な定数であり、ここでは2.0と定める。また、このような定数とせずに入力レベル lev の分散に比例した量とすることも考えられる。

【0088】そして、パラメータ生成部8は、出力端子105からidVUVパラメータを出力する。

【0089】このようにして、入力信号判定装置21は、最小レベル gml を更新しながら、それを元に音声信号区間と背景雑音区間とのしきい値(スレシヨルド) $\text{gml} \times B$ を変化できるので、このスレシヨルド $\text{gml} \times B$ とリファレンスレベル refLev を比較することにより音声信号区間と背景雑音区間との高精度な区別を可能とする。

【0090】ここで、図1に戻る。音声復号化装置31は、他の携帯電話装置の上記音声符号化装置20により可変レート符号化された上記符号化データをアンテナ26、アンテナ共用器25、受信機27、復調器29及び伝送路復号化器30を介して受け取り、復号化する。

【0091】この音声復号化装置31は、上述したように、音声符号化装置20から上記所定時間中にも常に伝送されてくるidVUV判定パラメータに基づいて、上記符号化データを復号化する。特に、idVUV判定パラメータが背景雑音区間を示す“1”であれば、8フレーム分において伝送されてきた上記複数種類のパラメータ、例えば雑音符号帳のシェイプインデックスや、ゲインインデックス、又はLSPパラメータを用いて、背景雑音を生成すると共に、8フレーム中では過去に送られてきた線スペクトル対(LSP)パラメータを補間して上記背景雑音を生成する。

【0092】実際に、音声復号化装置31では、常に、前回送られたLSP(prevLsp1)と前々回送られたLSP(prevLsp2)を、例えばRAM内に保持している。

【0093】そして、上記idVUV判定パラメータが背景雑音モードに入ると、新たなLSPは送られてこないの

で、 prevLsp1 、 prevLsp2 の更新を行わず、この二つのLSPを線形補間することにより、現在のフレームのLSPとし、背景雑音を形成する。

【0094】背景雑音モード中、8フレーム目に通常のUVとして音声符号化装置側からUVの全パラメータが送られてくるが、このときゲイン回路126からのゲインインデックスを調べ、インデックスが前回送られたインデックス+2より小さければ、そのフレームの合成に用いるLSPを前回送られたパラメータに置き換える。この動作については後述する。ただし、ゲインインデックスは小さい順にソートされているものとする。

【0095】このような音声復号化装置31の構成を図7及び図8に示す。図7は、音声復号化装置31の基本構成を示すブロック図である。

【0096】この図7において、入力端子202には上記図2の出力端子102からの上記LSP(線スペクトル対)の量子化出力としてのコードブックインデックスが入力される。入力端子203、204、及び205には、上記図2の各出力端子103、104、及び105からの各出力、すなわちエンベロープ量子化出力としてのインデックス、ピッチ、及びV/U/V判定出力がそれぞれ入力される。また、入力端子207には、上記図2の出力端子107からのUV(無声音)用のデータとしてのインデックスが入力される。

【0097】入力端子203からのエンベロープ量子化出力としてのインデックスは、逆ベクトル量子化器212に送られて逆ベクトル量子化され、LPC残差のスペクトルエンベロープが求められて有声音合成部211に送られる。有声音合成部211は、サイン波合成により有声音部分のLPC(線形予測符号化)残差を合成するものであり、この有声音合成部211には入力端子204及び205からのピッチ及びidVUV判定パラメータも供給されている。有声音合成部211からの有声音のLPC残差は、LPC合成フィルタ214に送られる。また、入力端子207からのUVデータのインデックスは、無声音合成部220に送られて、雑音符号帳を参照することにより無声音部分のLPC残差が取り出される。このLPC残差もLPC合成フィルタ214に送られる。LPC合成フィルタ214では、上記有声音部分のLPC残差と無声音部分のLPC残差とがそれぞれ独立に、LPC合成処理が施される。あるいは、有声音部分のLPC残差と無声音部分のLPC残差とが加算されたものに対してLPC合成処理を施すようにしてもよい。ここで入力端子202からのLSPのインデックスは、LPCパラメータ再生部213に送られて、LPCの α パラメータが取り出され、これがLPC合成フィルタ214に送られる。LPC合成フィルタ214によりLPC合成されて得られた音声信号は、出力端子201より取り出される。

【0098】ここで、入力端子205に供給されたidVU

V判定パラメータと入力端子207に供給された上記UVデータとしての雑音符号帳のシェイプインデックス及びゲインインデックスは、上記LPCパラメータ生成部213でのLPCパラメータの再生を制御するLPCパラメータ再生制御部240に送られる。

【0099】このLPC再生制御部240により制御され、LPCパラメータ再生部213は、背景雑音信号生成用のLPCを生成し、LPC合成フィルタ214に送る。

【0100】次に、図8は、上記図7に示した音声復号化装置31のより具体的な構成を示している。この図8において、上記図7の各部と対応する部分には、同じ指示符号を付している。

【0101】この図8において、入力端子202には、上記図2、3の出力端子102からの出力に相当するLSPのベクトル量子化出力、いわゆるコードブックのインデックスが供給されている。

【0102】このLSPのインデックスは、LPCパラメータ再生部213のLSPの逆ベクトル量子化器231に送られてLSP（線スペクトル対）データに逆ベクトル量子化され、スイッチ243を介してLSP補間回路232、233に送られてLSPの補間処理が施された後、LSP→ α 変換回路234、235でLPC（線形予測符号）の α パラメータに変換され、この α パラメータがLPC合成フィルタ214に送られる。ここで、LSP補間回路232及びLSP→ α 変換回路234は有声音（V）用であり、LSP補間回路233及びLSP→ α 変換回路235は無声音（UV）用である。またLPC合成フィルタ214は、有声音部分のLPC合成フィルタ236と、無声音部分のLPC合成フィルタ237とを分離している。すなわち、有声音部分と無声音部分とでLPCの係数補間を独立に行うようにして、有声音から無声音への遷移部や、無声音から有声音への遷移部で、全く性質の異なるLSP同士を補間することによる悪影響を防止している。

【0103】また、図8の入力端子203には、上記図2、図3のエンコード側の端子103からの出力に対応するスペクトルエンベロープ（Am）の重み付けベクトル量子化されたコードインデクスデータが供給され、入力端子204には、上記図2、図3の端子104からのピッチのデータが供給され、入力端子205には、上記図2、図3の端子105からのidVUV判定パラメータが供給されている。

【0104】入力端子203からのスペクトルエンベロープAmのベクトル量子化されたインデクスデータは、逆ベクトル量子化器212に送られて逆ベクトル量子化が施され、上記データ数変換に対応する逆変換が施されて、スペクトルエンベロープのデータとなって、有声音合成部211のサイン波合成回路215に送られている。

【0105】なお、エンコード時にスペクトルのベクトル量子化に先だってフレーム間差分をとっている場合には、ここでの逆ベクトル量子化後にフレーム間差分の復号を行ってからデータ数変換を行い、スペクトルエンベロープのデータを得る。

【0106】サイン波合成回路215には、入力端子204からのピッチ及び入力端子205からの上記idVUV判定パラメータが供給されている。サイン波合成回路215からは、上述した図2、図3のLPC逆フィルタ111からの出力に相当するLPC残差データが取り出され、これが加算器218に送られている。このサイン波合成の具体的な手法については、例えば本件出願人が先に提案した、特願平4-91422号の明細書及び図面、あるいは特願平6-198451号の明細書及び図面に開示されている。

【0107】また、逆ベクトル量子化器212からのエンベロープのデータと、入力端子204、205からのピッチ、idVUV判定パラメータとは、有声音（V）部分のノイズ加算のためのノイズ合成回路216に送られている。このノイズ合成回路216からの出力は、重み付き重畳加算回路217を介して加算器218に送っている。これは、サイン波合成によって有声音のLPC合成フィルタへの入力となるエクサイテイション（Excitation：励起、励振）を作ると、男声等の低いピッチの音で鼻づまり感がある点、及びV（有声音）とUV（無声音）とで音質が急激に変化し不自然に感じる場合がある点を考慮し、有声音部分のLPC合成フィルタ入力すなわちエクサイテイションについて、音声符号化データに基づくパラメータ、例えばピッチ、スペクトルエンベロープ振幅、フレーム内の最大振幅、残差信号のレベル等を考慮したノイズをLPC残差信号の有声音部分に加えているものである。

【0108】加算器218からの加算出力は、LPC合成フィルタ214の有声音用の合成フィルタ236に送られてLPCの合成処理が施されることにより時間波形データとなり、さらに有声音用ポストフィルタ238vでフィルタ処理された後、加算器239に送られる。

【0109】次に、図8の入力端子207s及び207gには、上記図3の出力端子107s及び107gからのUVデータとしてのシェイプインデックス及びゲインインデックスがそれぞれ供給され、無声音合成部220に送られている。端子207sからのシェイプインデックスは、無声音合成部220の雑音符号帳221に、端子207gからのゲインインデックスはゲイン回路222にそれぞれ送られている。雑音符号帳221から読み出された代表値出力は、無声音のLPC残差に相当するノイズ信号成分であり、これがゲイン回路222で所定のゲインの振幅となり、窓かけ回路223に送られて、上記有声音部分とのつながりを円滑化するための窓かけ処理が施される。

【0110】窓かけ回路223からの出力は、無声音合成部220からの出力として、LPC合成フィルタ214のUV（無声音）用の合成フィルタ237に送られる。合成フィルタ237では、LPC合成処理が施されることにより無声音部分の時間波形データとなり、この無声音部分の時間波形データは無声音用ポストフィルタ238uでフィルタ処理された後、加算器239に送られる。

【0111】加算器239では、有声音用ポストフィルタ238vからの有声音部分の時間波形信号と、無声音用ポストフィルタ238uからの無声音部分の時間波形データとが加算され、出力端子201より取り出される。

【0112】また、LPCパラメータ再生部213内部には、背景雑音信号生成用に用いられるLPCパラメータを再生するためのLSP補間回路245と、LSP→ α 変換回路247も上記スイッチ243の後段に設けられている。さらに、上記LSPの逆量子化器231により得られた上記prevLSP1とprevLSP2を保持しておくためのRAM244と、8フレームの間隔があいている上記prevLSP1とprevLSP2との補間をフレーム間隔に補正するためのフレーム補間部245も備えている。

【0113】LSP→ α 変換回路247からのBGN用の α パラメータは、LPC合成フィルタ214の無声音部分のLPC合成フィルタ237に送られる。

【0114】また、上記LPCパラメータ再生制御部240は、入力端子207gからの上記UVデータ用のゲインインデックスが、前回送られたインデックス+2より小さいか否かを判定するインデックス判定部242と、このインデックス判定部242からの判定結果と上記入力端子205から供給されるidVUV判定パラメータとに基づいて上記スイッチ243の切り換えを制御する切り換え制御部241とを備えてなる。

【0115】上記idVUV判定パラメータが1であるとき、すなわち背景雑音区間であることを示す上記基本パラメータとなるモードビットを受信したときのこの音声復号化装置31の動作を、図9に示すフローチャートを用いて説明する。

【0116】まず、ステップS21でidVUV判定パラメータが1であるとき、LPCパラメータ再生制御部240の切り換え制御部241は、切り換えスイッチ243をオフにする。そして、ステップS22に進み、LSP補間回路246で上記RAM244に保持されたPrevLSP1とPrevLSP2をフレーム補間回路245を通して得たフレーム毎の直線補間値を使ってBGN用のLSPを求める。そして、このBGN用のLSPは、UV用の合成フィルタ237に供給され、背景雑音が合成される。

【0117】なお、音声符号化装置20からは、背景雑音区間と判断されたフレームが連続8フレームとなると、次の9フレーム目は通常のUVデータ用のシェイプ

インデックス、ゲインインデックス及びLSPパラメータが送られてくる。ここで、音声符号化装置20では、9フレーム目に本当にidVUV判定パラメータがUVになることがないとは限らない。そこで、音声復号化装置31側では、本当のUV用データなのか、あるいは単に9フレーム目に送られたUV用の全パラメータなのかを判断する必要がある。

【0118】そこで、LPCパラメータ再生制御部240では、ステップS23でidVUV=0であるとき、ステップS24に進み、インデックス判定部242により入力端子207gを介して送られてきたUVデータ用のゲインインデックスを調べ、ステップS26のルーティンの処理か、ステップS27の本来のUVの処理かを判断する。具体的には、上述したように、入力端子207gからの上記UVデータ用のゲインインデックスが、前回送られたインデックス+2より小さいか否かを判定する。9フレーム目に送られてきたUVデータ用のインデックスが音声としてのUV用インデックスであれば前回送られたゲインインデックスのパラメータよりも大きいはずである。

【0119】ステップS24で上記ゲインインデックスが前回送られてきたインデックス+2より小さいとなれば、これは9フレーム目に送られたUVデータ用の、すなわちルーティン用のデータであると判断し、ステップS26に進み、切り換え制御部241により、スイッチ243をLSP補間回路246側に接続し、上記PrevLSP1とPrevLSP2を用いた直線補間により求めた値に変えて、UV用に送られたLSP逆量子化部231からのLSPパラメータをLSP補間回路246に供給する。LSP補間回路246では、このアップデートされたLSPパラメータを、そのままLSP→ α 変換回路247に供給する。そして、LSP→ α 変換回路247からのBGN用の α パラメータは、LPC合成フィルタ214の無声音部分のLPC合成フィルタ237に送られ、9フレーム目には8フレーム間の背景雑音とは異なった背景雑音が得られることになる。このため、背景雑音の不自然さを緩和することができる。

【0120】また、ステップS24で上記ゲインインデックスが前回送られてきたインデックス+2以上より大きいとなれば、これは9フレーム目に送られてきたのは、本当のUV用のパラメータであると判断し、ステップS27に進む。ステップS27では、切り換え制御部241がスイッチ243をUV用のLSP補間回路233に切り換えて、通常のUV用のLSP補間により得られたLSPを使った無声音の合成が行われる。

【0121】一方、上記ステップS23でidVUV判定パラメータが0でないと判断すると、ステップS25に進み、UV用のLSPから α パラメータを変換し、合成フィルタ236で有声音を合成する。

【0122】以上のように、音声復号化装置31では、idVUV判定パラメータが1であるとき、すなわち他の携

帯電話装置の音声符号化装置が背景雑音区間を検出したときには、8フレームにおいて伝送されてきた複数種類のパラメータを用いて上記背景雑音を生成すると共に、上記8フレーム中には過去に送られたパラメータを用いて上記背景雑音を生成する。このため、背景雑音の不自然さを緩和することができる。また、9フレーム目に偶然に、本当の無声音区間が検出されて伝送されてきても、正確に判断できるので、高品質な音声を復号できる。

【0123】LSPは8フレーム分のディレイが生じることになるが、背景雑音モード中は完全に滑らかに繋がることになり、急激にLSPが変化して異音が発生することがなくなる。また、音声の子音部は背景雑音より高いレベルを有することが多いので、音声の子音部を誤って背景雑音として処理してしまうことを防げる。

【0124】そして、再び背景雑音モードに入ったらprevLsp1、prevLsp2を線形補間することにより、現在のフレームのLSPとする。

【0125】ところで、モードビットが充分にあり、これを一つのモードとして送ることができれば、このような処理は必要ない。また、prevLsp1とprevLsp2の補間方法は線形補間以外にも様々なものが考えられる。

【0126】

【発明の効果】本発明に係る音声符号化方法及び装置は、可変レート符号化を効率良く実現する。

【0127】また、本発明に係る音声復号化方法及び装置は、可変レート符号化を実現する音声符号化方法及び装置により符号化された符号化データを用いて、背景雑音を不自然感を緩和して生成できる。

【図面の簡単な説明】

【図1】本発明に係る音声符号化方法及び装置、並びに音声復号化方法及び装置の実施の形態となる携帯電話装置の構成を示すブロック図である。

【図2】上記携帯電話装置を構成する音声符号化装置の基本的な構成を示すブロック図である。

【図3】上記図2に示した音声符号化装置の詳細な構成を示すブロック図である。

【図4】上記音声符号化装置内部にあって入力信号を判定する入力信号判定装置の構成を示すブロック図である。

【図5】上記図4に示した入力信号判定装置を構成する最小レベル演算部のアルゴリズムを説明するためのフローチャートである。

【図6】上記図4に示した入力信号判定装置を構成するリファレンスレベル演算部を説明するための特性図である。

【図7】上記携帯電話装置を構成する音声復号化装置の基本的な構成を示すブロック図である。

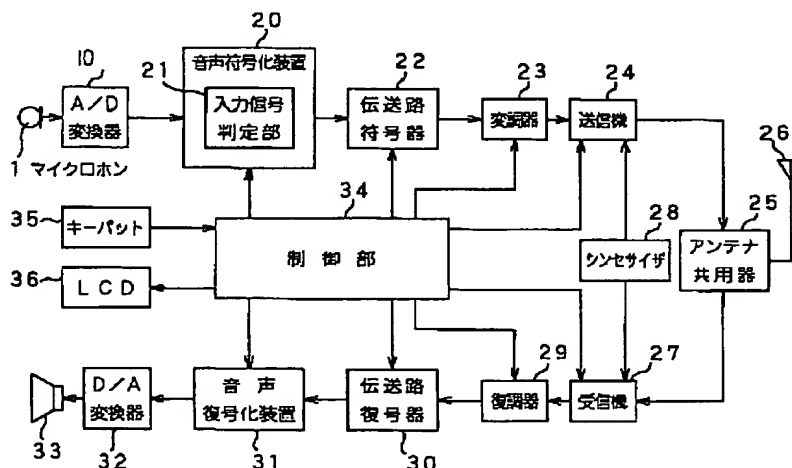
【図8】上記図7に示した音声復号化装置の詳細な構成を示すブロック図である。

【図9】上記音声復号化装置の動作を説明するためのフローチャートである。

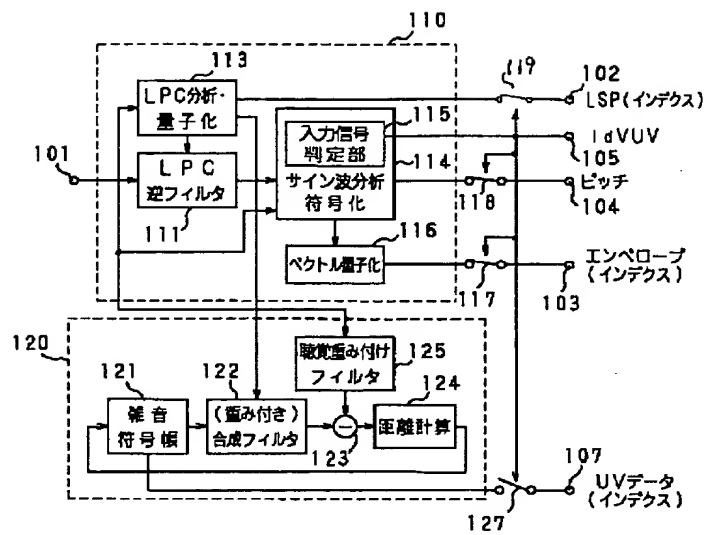
【符号の説明】

20 音声符号化装置、21、115 入力信号判定部、31 音声復号化装置、240 LPCパラメータ再生制御部、241 切り換え制御部、242 インデックス判定部

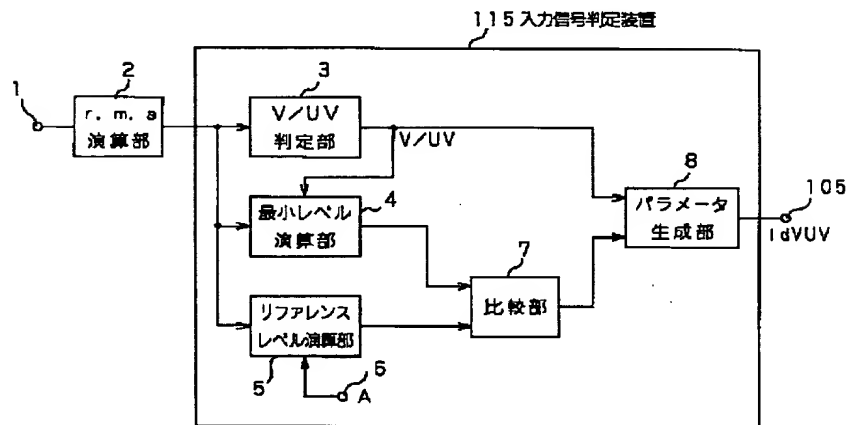
【図1】



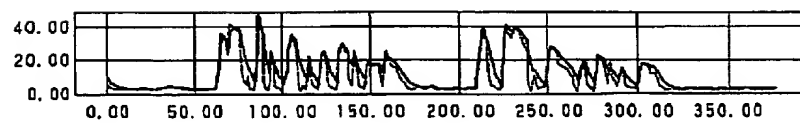
【図2】



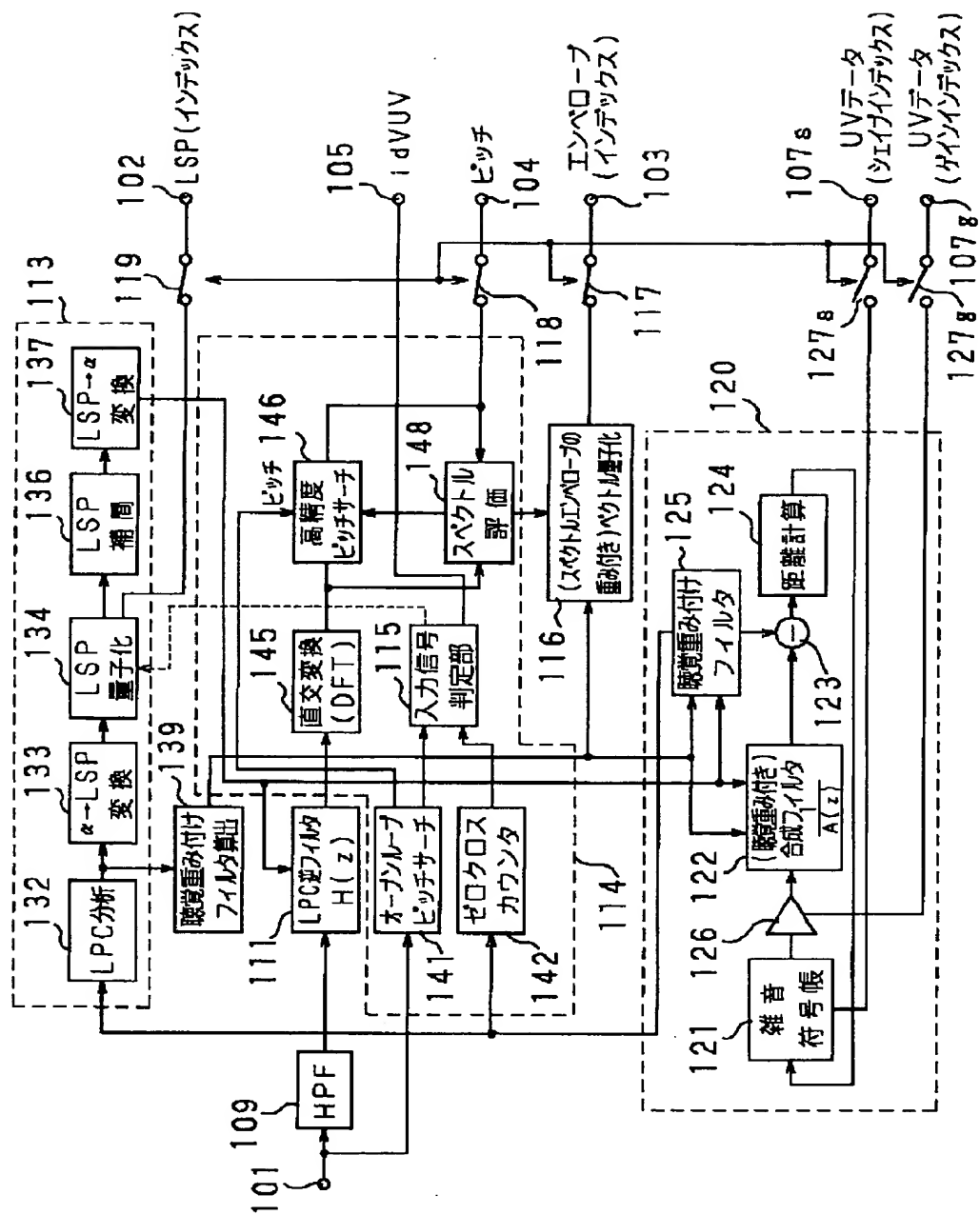
【図4】



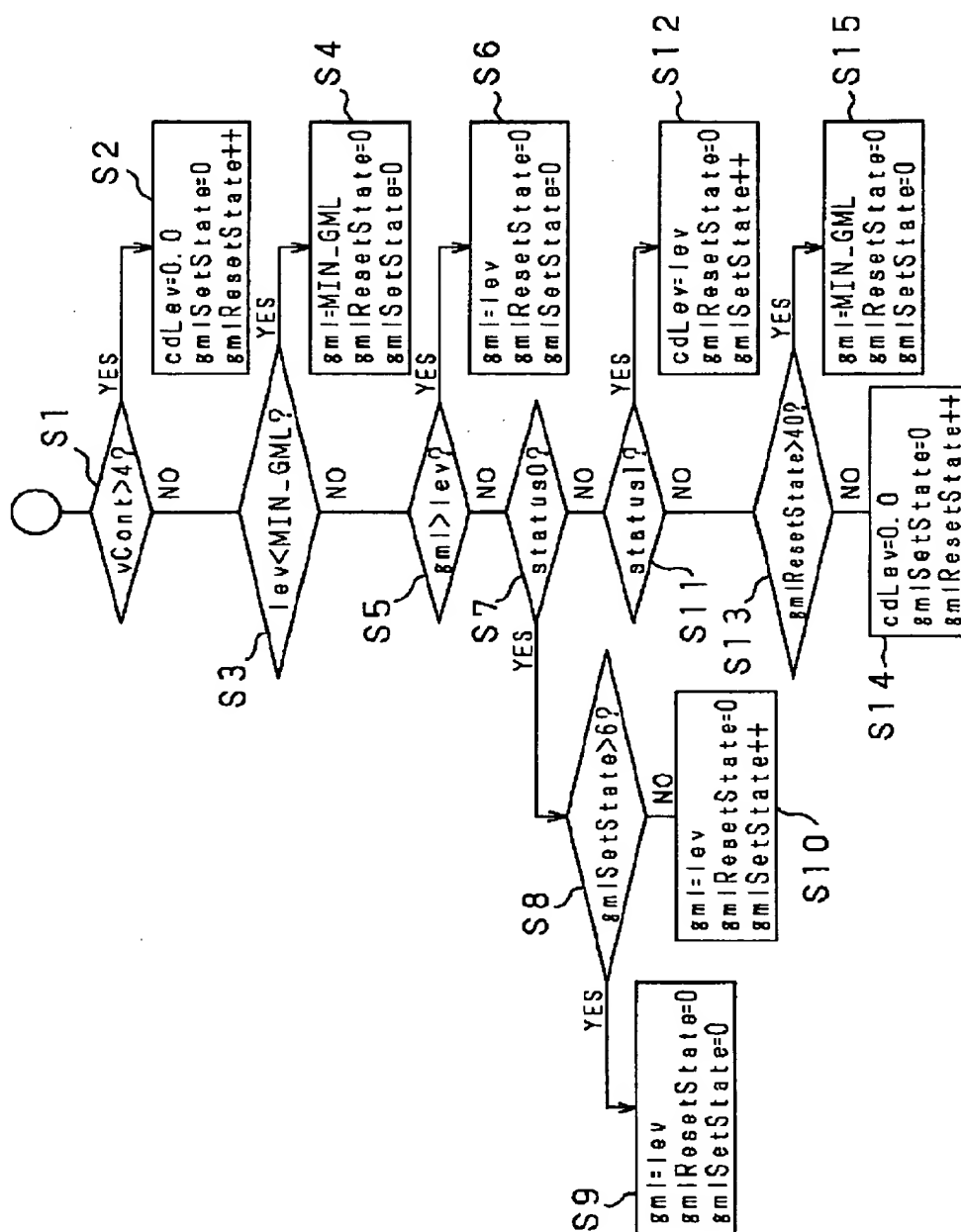
【図6】



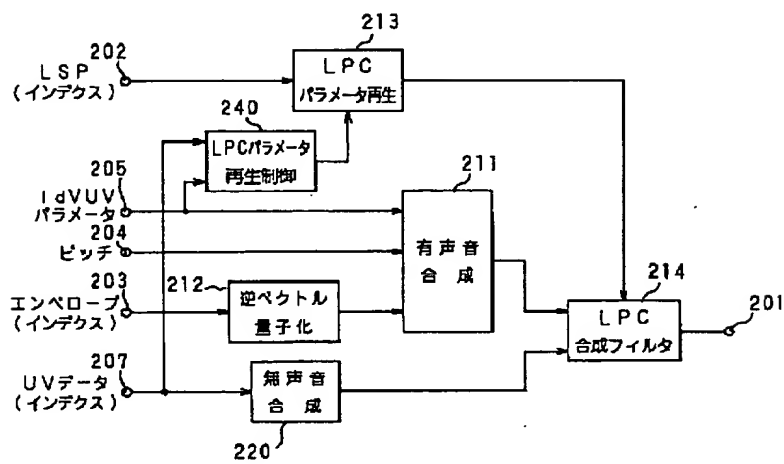
【図 3】



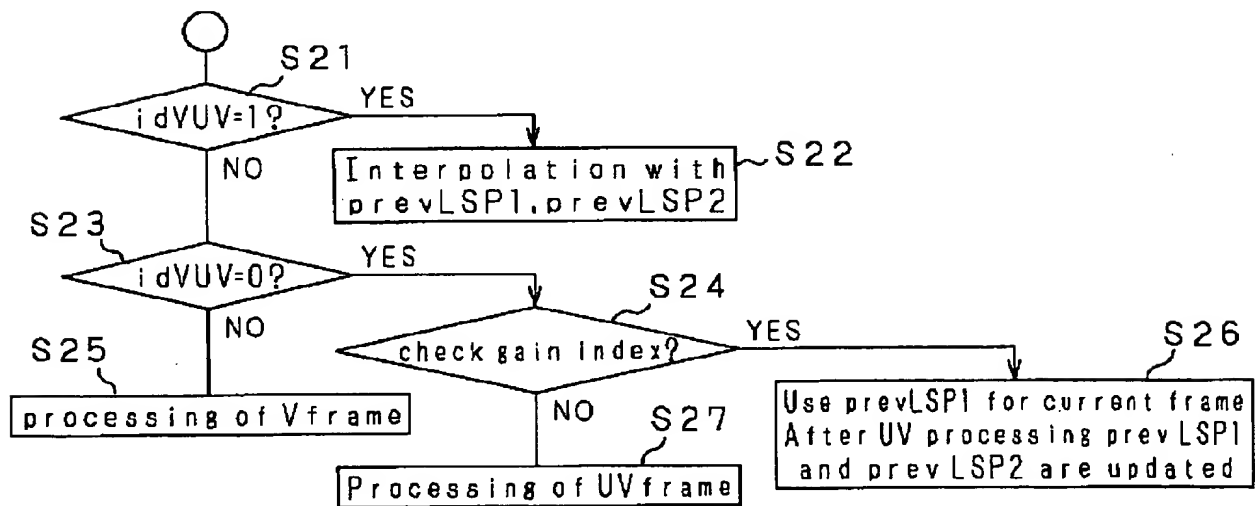
【図5】



【図7】



【図9】



【图 8】

